



Data Sharing and Intellectual Capital Working Group Teleconference

June 10, 2004	2:00 pm EDT
Attendees:	<p><i>Attendees:</i></p> <p>City of Hope: Joyce Niland Fred Hutchinson: Mark Thornquist Jackson Laboratory: Barbara Tennent Oregon Health and Science University: Edwin Quick Thomas Jefferson University—Kimmel: Jack London University of Arizona: David Mount University of Iowa: Terry Braun; Thomas Cassavant University of Pittsburgh: Michael Becich Washington University—Siteman: Mark Watson U Penn – Howard Bilofsky Fox Chase – Amin Chisti; Pat Harsche-Weeks NCI - Wendy Patterson; Leslie Derr BAH - Phan Winter; Mike Keller</p>
Introduction	<p>Wendy Patterson opened the meeting, reviewed the agenda, and asked whether the group had comments on the notes from the 5/27/04 teleconference. Since no comments were received, participants were encouraged to send any comments to Phan who will post the notes in final form.</p>
Report from Liaisons	<p><u>Training:</u> Ed Quick (OHSU)</p> <p>The most recent Training WG meeting took place on June 2. Three Special Interest Groups (SIGs) for Developers, Adopters, and Communication have been formed. The Training Working Group discussed contract matters and began the process of evaluating caBIG webcast capabilities and conferencing options. The Training WG has delayed its face-to-face meeting originally scheduled in June pending negotiation of contracts. The NCI's Center for Strategic Dissemination is in the process of developing a needs assessment for caBIG Cancer Centers regarding training materials. The Adopter SIG raised the idea of partnering with</p>



commercial groups to develop training materials, and will further explore this option.

Strategic Planning: Michael Becich (Pittsburgh)

The Strategic Plan WG is in the process of developing three major recommendations, which will:

- (1) request that the Architecture Workspace draft a document articulating the principles underlying the caBIG initiative ASAP;
- (2) request that the Vocabulary Workspace develop and rationalize Common Data Element libraries so that collaborative work can begin; and
- (3) assemble a subcommittee of the Strategic Planning WG to identify a list of potential pitfalls ("Gotcha's") for caBIG. The list of "Gotcha's" will include political, sociological, and technical issues, such as intellectual property restrictions and management of expectations that must be identified early on to avoid the derailment of caBIG. The issues will be assigned to specific groups with immediate expertise but other Workspaces and Working Groups will be asked to weigh in as required.

These recommendations will be shared with other WS and WG members for comment and action.

Integrative Cancer Research: Terry Braun (University of Iowa – Holden)

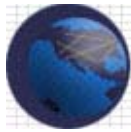
There were two presentations during the last ICR meeting, so there was not much time for discussion. The first presentation concerned the development and use of Common Data Elements (CDE); the second presentation provided an overview of the caBIG CDE Development Governance Model.

Clinical Trials: Don Connelly (University of Minnesota)

Don Connelly was absent so there was no report.

Tissue Banks and Pathology Tools Workspace: Mark Watson (Washington University - Siteman)

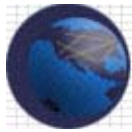
The TBPT WS is developing a system to gather information



	<p>from adopters with the goal of capturing institutional workflow for pathology laboratories. The Requirements Specification Survey was distributed to all TBPT WS participants and completed surveys were due by Friday, June 11.</p>
Report on IP POCs	<p><u>Iowa</u>: Daniel G. Happe, J.D. <u>City of Hope</u>: Larry Couture, Ph.D. <u>Washington University – Siteman</u>: Jon Kratochvil</p> <p>DSIC WG members were again encouraged to send in POC information to Phan.</p>
Report on Questionnaire Development	<p>Wendy noted the revisions to the questionnaire from the previous version and asked for comments.</p> <p>Pat Harsche-Weeks thought that the introductory section was very good. She thought that the survey was comprehensive but she advised the group to be aware that each responding institution may have to review hundreds of existing agreements and that searching for these records will not be a trivial task. The group needs to be realistic about the time required to respond to the questionnaire. For example, Pat thinks that this effort could take up to 2 months at Fox Chase. Pat also urged the group to send the survey to persons responsible for reviewing clinical trial agreements as well as to technology transfer offices.</p> <p>Mark Watson thought that responses would vary in terms of response time and content, which should be a good indication of how well the responding participants have organized and maintained their intellectual property departments.</p> <p>Wendy asked whether there were any opinions as to who should send the survey. The group suggested that Booz Allen send it to the DSIC WG IP POCs with copies to WG members and other appropriate parties.</p> <p>Howard Bilofsky reiterated the importance of recognizing the level of effort required for meaningful survey responses effort and asked whether there would be funding for this activity. Wendy agreed to look into this and follow up.</p> <p>Other suggestions were to include:</p> <p>(1) clinical trial agreements and agreements providing</p>



	<p>access to software and/or algorithms in the introductory section's list of agreements; and</p> <p>(2) academic-academic collaborations in the section on existing agreements.</p> <p>Wendy asked whether the caBIG online forum is a good format for providing comments to the survey. The group responded that they generally prefer responding to documents transmitted by e-mail. Wendy will update the current version of the questionnaire and Phan will send it to the group by e-mail.</p> <p>The next step is to submit the revised survey to DSIC WG IP POCs to get a realistic estimate on the time required to respond. The group reiterated the importance of circulating the draft questionnaire to industry representatives; Ardais and Glaxo Smith Kline were suggested as candidate reviewers. The group did not think it was realistic to obtain industry input before the fall.</p> <p>Wendy concluded by stating that it was important for the group to identify its IP POCs so that their input on the questionnaire can be solicited.</p>
Review of Data Access Hierarchy	<p>The group then resumed discussion of the data hierarchy first articulated during the 4/29/04 teleconference:</p> <ul style="list-style-type: none">• No sharing at all• Sharing restricted to active collaborators• Limited sharing within a broader research community (e.g., the NCI's Early Detection Research Network (EDRN))• Open data sharing without any restrictions (e.g., data are accessible through a website) <p>Howard Bilofsky thought that the tiers of access were more complex than listed in that data access is often determined by a researcher's role and/or function. He agreed to refine the data access hierarchy and send it to Phan before the next teleconference.</p> <p>Mark Thornquist thought that the hierarchy should be more specific as to individual data elements. He noted that within any given data set, different levels of restrictions are frequently imposed on individual data elements. Consequently, researchers may not want to provide unrestricted access to</p>



complete data sets to all requestors. Along these lines, Mark Watson pointed to practices surrounding the use of microarray data sets. In this context, various researchers are granted different levels of access. He thought that the DSIC WG could develop examples of different levels of sharing by looking at these group-share sites. Ed Quick commented that as the grid grows, it is very important to capture the origins of various data points (i.e., the data's provenance) so that future users will know how to obtain access.

The group agreed that it is important to understand the practical constraints with respect to sharing confronted by researchers participating in caBIG. The discussion moved to the implication of announcing the existence of a dataset. The group considered whether the announcement that certain sets of data exist (or are in the process of being generated) obligate the data creator to share the data within a certain period of time. It was acknowledged that the hierarchy creates a timeline and that refusal to provide any indication of when data would be shared would not fit within the caBIG philosophy. However, it was generally agreed that data sharing standards that prescribe a set time period for data release would not be practical. Howard thought that a release date would emerge by consensus. Pat Harsche-Weeks suggested that researchers be asked to state when their data would be available.

The group conjectured that the ownership culture needs to change, such that time-related constraints on sharing are negotiated. One suggested approach was to place the obligation on the person controlling the data to inform future recipients when data will be available.

Wendy pointed out that there seems to be a "we vs. they" distinction in sharing discussions, but she is not sure who "they" are. She suggested viewing the discussion of these issues within the context of the caBIG community. In response, members of the group stated that the success of caBIG will depend on whether researchers believe that benefits derive from participating in caBIG. The research community needs to be assured that they receive value in exchange for sharing its data. Currently, the value of data is achieved through publication, which can lead to tenure. In order to promote data sharing, caBIG needs to change the culture of data ownership to assign credit for sharing. Ed Quick noted that the Communication SIG of the Training Working Group is



Data Sharing and Intellectual Capital Working Group Teleconference, June 10th, 2004, 2:00 pm EDT

	also working on the issue of perceived benefits of participating in caBIG and recommended that we work more closely with that group on this issue.			
Items for Next Teleconference: June 24th	Wrap up discussion of survey and send to IP POCs. Continue discussion of data sharing hierarchy.			
Action Items:				
	Name Responsible	Action Item	Date Due	Notes
	Howard Bilofsky	Refine sharing hierarchy	6/23/2004	Will send new language to Phan
	Wendy Patterson	Revise questionnaire	6/21/2004	
	Phan Winter	Circulate revised questionnaire	6/21/2004	
	All	Send IP POC to Phan	6/4/2004	Phan will call Centers to solicit POC
	Phan Winter	Solicit IP POC	6/21/2004	